8-2003

The Parasitome of the Phytonematode Heterodera glycines

Bingli Gao
University of Georgia

R. Allen
University of Georgia

Tom Maier
Iowa State University, trmaier@iastate.edu

Eric L. Davis
North Carolina State University

Thomas J. Baum
Iowa State University, tbaum@iastate.edu

See next page for additional authors

MPMI Vol. 16, No. 8, 2003, pp. 720–726. Publication no. M-2003-0515-01R. © 2003 The American Phytopathological Society

The Parasitome of the Phytonematode *Heterodera glycines*

Bingli Gao,[1] R. Allen,[1] Tom Maier,[2] Eric L. Davis,[3] Thomas J. Baum,[2] and Richard S. Hussey[1]

[1]Department of Plant Pathology, University of Georgia, Athens 30602-7274, U.S.A.; [2]Department of Plant Pathology, Iowa State University, 351 Bessey Hall, Ames 50011, U.S.A.; [3]Department of Plant Pathology, North Carolina State University, Box 7616, Raleigh 27695-7616, U.S.A.

Submitted 12 February 2003. Accepted 14 April 2003.

Parasitism genes expressed in the esophageal gland cells of phytonematodes encode secretions that control the complex process of plant parasitism. In the soybean cyst nematode, *Heterodera glycines*, the parasitome, i.e., the secreted products of parasitism genes, facilitate nematode migration in soybean roots and mediate the modification of root cells into elaborate feeding cells required to support the growth and development of the nematode. With very few exceptions, the identities of these secretions are unknown, and the mechanisms of cyst nematode parasitism, therefore, remain obscure. The most direct and efficient approach for cloning parasitism genes and rapidly advancing our understanding of the molecular interactions during nematode parasitism of plants is to create gland cell–specific cDNA libraries using cytoplasm microaspirated from the esophageal gland cells of various parasitic stages. By combining expressed sequence tag analysis of a gland cell cDNA library with high throughput in situ expression localization of clones encoding secretory proteins, we obtained the first comprehensive parasitome profile for a parasitic nematode. We identified 51 new *H. glycines* gland-expressed candidate parasitism genes, of which 38 genes constitute completely novel sequences. Individual parasitome members showed distinct gland cell expression patterns throughout the parasitic cycle. The parasitome complexity discovered paints a more elaborate picture of host cellular events under specific control by the nematode parasite than previously hypothesized.

Additional keywords: microaspiration, plant-parasitic nematode.

Secretions from parasites and pathogens are the primary functional and signal molecules at the host interface (Davis et al. 2000; Hussey 1989; Petnicki-Ocwieja et al. 2002). An intricate exchange of molecular signals and responses between parasite and host occurs at this interface that changes in both temporal and spatial balance as the developmental needs of both "players" change during the interaction. Often, the true nature of these interactions can only be observed in vivo, especially in the case of obligate parasites such as phytonematodes. The majority of phytonematodes are parasites of plant roots, including the soybean cyst nematode *Heterodera glycines*, which causes an estimated $1 billion in U.S.A. soybean losses annually (Davis and Tylka 2000). Identifying vulnerable points in the parasitic cycle of *H. glycines* to target for disruption offers great promise to design management tactics to limit soybean damage due to this pathogen.

Understanding the complexity of the molecular signal exchange and response during host-parasite interactions has remained a daunting task, until recent advances in molecular biology and genomic science have allowed analyses of global gene expression on scales as small as a single cell (De Boer et al. 2002a; Karrer et al. 1995; Rekhter and Chen 2001). For single-cell organisms such as pathogenic bacteria, analyses of gene products destined for secretion during interaction with the host can be analyzed on a "whole organism" scale (Petnicki-Ocwieja et al. 2002). However, in multicellular parasites such as phytonematodes, whole organism analyses of secreted gene products may not provide specific information on molecules directly involved in parasitism.

The most evolutionarily advanced adaptations for plant parasitism by nematodes are the products of parasitism genes expressed in their esophageal gland cells and secreted through their stylet (feeding spear) into host tissue and cells (Davis et al. 2000; Hussey 1989). Stylet secretions mediate nematode infection and parasitism of plants, and developmental changes in the secreted proteins occur during the parasitic cycle. The secreted products of the parasitism genes expressed in the nematode's esophageal gland cells constitute in part or in total the "parasitome", i.e., a subset of the secretome (secreted proteins) of a parasite that mediates parasitism (based on the nomenclature of Greenbaum et al. 2001; Hussey et al. 2002). For phytonematodes the definition of parasitome is restricted in the narrow sense to secretions released through the stylet, because host plant cell modifications induced by these nematodes appear to be limited to cells coming in direct contact with their stylet (Hussey and Grundler 1998). The first members of a phytonematode parasitome to be cloned were β-1,4-endoglucanase (cellulase) genes expressed in the two subventral esophageal gland cells of cyst nematodes (Smant et al. 1998; Yan et al. 1998). These β-1,4-endoglucanase genes are developmentally expressed in the subventral gland cells in *H. glycines* (De Boer et al. 1999), and a β-1,4-endoglucanase has been determined to be secreted within soybean roots by parasitic juveniles during the intracellular migration phase of the infection process, confirming the in planta secretion of a parasitism gene product (Wang et al. 1999).

Differential gene expression has been the most widely used method to clone parasitism genes expressed in the esophageal gland cells of phytonematodes but with only limited success (Ding et al. 1998; Lambert et al. 1999; Qin et al. 2000). Recently, a direct and efficient cell-specific method for cloning

Corresponding author: R. S. Hussey: Telephone: +1-706-542-1254; Fax: +1-706- 542-1262; E-mail: hussey@arches.uga.edu

Nucleotide and amino acid sequence data have been submitted to the DDBJ/EMBL/GenBank databases.

nematode parasitism genes was developed by microaspirating the cytoplasm from the esophageal gland cells (two subventral and one dorsal) of a range of *H. glycines* parasitic stages to provide mRNA for construction of gland-cell cDNA libraries using reverse-transcriptase polymerase chain reaction (RT-PCR) (Gao et al. 2001a; Wang et al. 2001). Both suppression subtractive hybridization (Gao et al. 2001a) and signal peptide selection (Wang et al. 2001) of gland-cell cDNA libraries have provided a sampling of parasitism genes expressed within *H. glycines*, but these data suggested that a more comprehensive approach was necessary to obtain a complete profile of the nematode parasitome. As illustrated below, coupling expressed sequence tag (EST) analyses with techniques designed to isolate and characterize genes expressed exclusively within the esophageal gland cells of *H. glycines* has provided, for the first time, direct and surprising information on the complexity and dynamics of the parasitome of a multicellular parasite.

RESULTS

Analysis of a *H. glycines* gland library.

The direct method of microaspirating the cytoplasm from esophageal gland cells of only 10 parasitic stages of *H. glycines* provided sufficient mRNA to construct a gland-cell cDNA library representative of parasitism gene expression within the gland cells throughout the parasitic cycle. Macroarraying *H. glycines* gland cell cDNA clones on a nylon membrane for indexing and subtraction with an intestinal cDNA library promoted efficient identification of unique candidate parasitism genes. High quality sequences were obtained for 3,711 sequences, which produced 716 contigs representing 2,198 sequences (leaving 1,513 singletons). The deduced protein sequences of 261 cDNAs from 3,711 cDNA sequences were predicted by SignalP to contain amino-terminal signal peptides to direct the proteins into the secretory pathway. The presence of the predicted signal peptides identified these esophageal gland cell proteins as candidates for being secreted by the nematode and potentially having a biological function in the *H. glycines*–soybean interaction. These clones were selected for in situ mRNA hybridization to confirm gland expression within nematode specimens. These analyses identified 53 distinct cDNA clones that specifically hybridized to transcripts accumulating within the subventral (12 clones) or dorsal (41 clones) esophageal gland cells of *H. glycines* (Table 1, Fig. 1). Fulfilling the two criteria of gland expression and encoding proteins with signal peptides (for potential secretion), these clones can be regarded as putative parasitism genes and their protein products as members of the *H. glycines* parasitome. Full-length cDNA sequences (from 5′ cap to 3′ polyA(+) tail) with predicted open reading frames were obtained for 51 candidate parasitism genes directly from the library. 5′–rapid amplification of cDNA ends (RACE) with a SMART race cDNA amplification kit (Clontech) was used according to manufacturer's instructions to obtain full-length cDNAs for two clones (13C08 and 45D07). The full-length cDNA sequences for the candidate parasitism genes ranged in size from 439 to 1,937 bp (Table 1). The predicted proteins of several of the putative parasitism gene contig members differed slightly in amino acid identity and, therefore, represented gene families, e.g., predicted proteins of clones 10A07, 20G04, and 27D09 shared 86 to 95% sequence identity (Table 1). Most interesting, PSI-BLASTP analyses revealed 75% of the putative parasitism genes to encode novel protein sequences specific to *H. glycines* ($E < 0.005$). While 62% of all predicted *H. glycines* gland cell proteins had homologues in the free-living nematode *Caenorhabditis elegans*, only 9% of the products of the putative parasitism genes were homologous with proteins of *C. elegans*. None of the 53 putative parasitism gene cDNAs contained the nematode-conserved 22-nt trans-spliced leader (SL1) or derivatives thereof (Blumenthal and Steward 1997; Ray et al. 1994).

Developmental expression.

Developmental expression patterns of the parasitism genes varied greatly throughout the *H. glycines* parasitic cycle (Table 1, Fig. 1). While mRNA of most parasitism genes expressed in the subventral gland cells was usually detectable only in the nonfeeding preparasitic and migratory parasitic second-stage juveniles (Fig. 1A), transcripts of a few subventral gland cell-expressed parasitism genes were also detected in feeding sedentary second-stage juveniles and later parasitic stages (Fig. 1B). On the other hand, parasitism gene expression in the dorsal gland cell was usually restricted to the feeding sedentary parasitic stages (Fig. 1C through F, H, and I), with transcripts of only a few genes detectable in the dorsal gland cell in nonfeeding preparasitic and migratory parasitic second-stage juveniles (Fig. 1G).

Secretory nuclear proteins.

Fifteen of the predicted proteins of the putative parasitism genes were predicted by PSORT II to contain canonical nuclear localization signals (NLS) and to be targeted to the nucleus (Table 1). mRNA of three predicted nuclear proteins (4E02, 10C02, and 17G06) was detectable exclusively within the two subventral gland cells (Fig. 1A and B), whereas mRNA of the remaining NLS-containing proteins was detectable only within the dorsal gland cell of *H. glycines*. Thirteen of the candidate secretory nuclear proteins were novel sequences with no significant homology with characterized proteins in the databases (Table 1). Two others, 8H07 and 10A06, had significant similarity to SKP-1 protein and a hypothetical rice protein, respectively.

Other proteins with homologues in other organisms.

Relatively few of the *H. glycines* parasitism gene products were similar to known proteins in other organisms. Four subventral gland–expressed genes were predicted to encode cell wall- and pectin-degrading enzymes; clones 6F06, 13C08, and 26D05 had significant similarity to β-1,4-endoglucanases from nematodes, and clone 17C07 encoded a pectate lyase. The predicted product of clone 3B05, expressed in the subventral gland cells in the parasitic stages, consisted of only a cellulose-binding domain and a signal peptide. Clone 2A05 encoded a venom allergen-like protein and clone 3D11 encoded a chitinase, both of which were expressed in the subventral gland cells. Clone 45B07, which was expressed in the dorsal gland cell of all *H. glycines* life stages examined, encoded a chorismate mutase, a branch point enzyme in the shikimate pathway in plants. The predicted product of clone 4F01 was similar to an annexin, also expressed in the dorsal gland cell of parasitic stages. Two other dorsal gland–expressed clones, 3H07 and 4G06, encoded secretory ubiquitin proteins.

DISCUSSION

Although it is currently not possible to predict the number of parasitism genes in *H. glycines*, we expect only a small fraction of the estimated 15,000 to 20,000 protein-coding genes (based on the approximately 19,000 protein-coding genes of *C. elegans* [The *C. elegans* Sequencing Consortium 1998]) of a phytonematode to have a direct role in parasitism. Previously, only a very few parasitism genes have been identified, which was accomplished mostly in studies comparing gene expression between specific nematode stages or tissue (Ding et al. 1998; Lambert et al. 1999; Qin et al. 2000). Therefore, a comprehensive analysis of the nematode parasitome requires a cell-

specific approach involving a more direct analysis of gene expression in the nematode's esophageal gland cells. Indeed, our discovery of an extensive suite of 53 putative parasitism genes, of which 51 were previously unidentified, validates our experimental strategy of directly targeting via microaspiration the gland cells of a range of parasitic stages coupled with EST analysis and in situ expression localization to obtain a comprehensive profile of the parasitome of *H. glycines*. Although some candidate parasitism gene products could possibly encode membrane proteins involved in the secretory pathway rather than proteins for actual secretion, none of the predicted products of the candidate parasitism genes contain transmembrane domains based on the TMHMM computer analysis used to predict transmembrane helices in proteins.

Table 1. Summary of 53 candidate parasitism genes encoding proteins preceded by a signal peptide for secretion and expressed exclusively within the esophageal gland cells of *Heterodera glycines*.

					Gland expression[d]		
Clone[a]	**Accession no.[b]**	**FL/ORF (bp)[c]**	**Highest protein similarity**	**BLASTP score/E value**	**Pre-J2**	**Par-J2**	**J3-A**
2A05	AY028639	683/439	MI-MSP-1—*Meloidogyne incognita*	114/1e^{-24}	SvG	SvG	SvG
2B10	AF273728	607/420	Gland cell protein—*H. glycines*	0	–[f]	DG	DG
2D01	AF469057	711/558	Pioneer[e]		–	DG	DG
3B05	AF469058	585/423	CBP - *H. glycines*	35/.19	–	SvG	SvG
3D11	AF468679	1120/10533	Chitinase - *Caenorhabditis elegans*	274/2.7e^{-21}	–	SvG	SvG
3H07	AF473831	571/318	Ubiquitin extension—*Nicotiana tobacco*	136/5e^{-32}	DG	DG	DG
4D06	AF469063	750/615	Pioneer		–	DG	DG
4D09	AF469061	738/501	Pioneer		DG	DG	DG
4E02*	AF473826	449/279	Pioneer		SvG	SvG	SvG
4F01	AF469059	1174/1026	Annexin—*C. elegans*	242/4e^{-63}	–	DG	DG
4G05	AF473830	928/765	Pioneer		–	DG	DG
4G06	AF469060	613/360	Hexaubiquitin—*Helianthus annuus* (85% identity to 3H07)[g]	151/1e^{-36}	DG	DG	DG
4G12	AF473827	621/417	Pioneer (91% identity to 2B10)		–	DG	DG
5D06*	AF469062	1937/1470	Pioneer		–	DG	DG
5D08*	AF473828	693/441	Pioneer		–	DG	DG
6E07*	AF473829	1046/645	Pioneer		–	DG	DG
6F06	AY043224	1333/1059	Cellulase—*H. glycines*	601/1e^{-170}	SvG	SvG	
7E05	AF500023	518/330	Pioneer		–	DG	DG
8H07*	AF500024	1457/1197	SKP1-like protein—*Arabidopsis thaliana*	94/3e^{-18}	–	DG	DG
10A06*	AF502391	1239/927	RING-H2 zinc finger protein—*A. thaliana*	50/3e^{-05}	–	DG	DG
10A07*	AF500021	837/729	Pioneer		–	DG	DG
10C02*	AF500017	449/279	Pioneer (92% identity to 4E02)		–	SvG	SvG
11A06	AF500015	673/561	Pioneer (91% identity to 2D01)		–	DG	DG
12H04	AF490244	1908/1614	Pioneer		DG	DG	DG
13A06*	AF500020	899/669	Pioneer (95% identity to 6E07)		–	DG	DG
13C08	AF469055	1101/1002	Cellulase—*H. glycines*	270/1e^{-71}	SvG	SvG	
16B09	AF490246	676/555	Pioneer		–	DG	DG
17C07	AF520566	957/792	Pectate lyase—*H. glycines*	461/e^{-129}	SvG	SvG	
17G06*	AF490247	600/300	Pioneer		–	SvG	
18H08	AF490248	632/399	Pioneer		–	DG	DG
19B10	AF490249	782/666	Pioneer		–	DG	DG
19C07	AF490250	660/333	Pioneer		–	DG	DG
20E03	AF490251	654/579	Pioneer		–	SvG	SvG
20G04*	AF500022	816/648	Pioneer (95% identity to 10A07)		–	DG	DG
21E12*	AF500028	439/354	Pioneer		–	DG	DG
22C12	AF500029	676/549	Pioneer (92% identity to 16B09)		–	DG	DG
23G12	AF500033	605/321	Pioneer		DG	DG	DG
24A12	AF500034	598/441	Pioneer		–	DG	DG
25A01	AF500019	750/528	Pioneer		–	DG	DG
25G01	AF006052	1600/1428	*Hg-eng-1—H. glycines*	0	SvG	SvG	
26D05	AY101191	1125/1008	Cellulase—*Pratylenchus penetrans*	263/2e^{-69}	SvG	SvG	SvG
27D09*	AY101190	851/708	Pioneer (86% identity to 10A07)		–	DG	DG
28B03	AF500025	1500/1302	Pioneer		DG	DG	DG
29D09	AF500016	757/615	Pioneer (95% identity to 4D06)		–	DG	DG
30C02	AF502393	537/492	Pioneer		–	DG	DG
30D08*	AF500027	443/384	Pioneer (82% identity to 21E12)		–	DG	DG
30E03	AF500035	675/558	Pioneer (98% identity to 16B09)		–	DG	DG
30G12	AF500018	881/717	Pioneer (93% identity to 4G05)		–	DG	DG
32E03*	AF500036	701/588	Pioneer		–	DG	DG
33A09	Ay125963	461/270	Pioneer		DG	DG	
33E05	AF502392	684**	Pioneer		DG	DG	
34B08	AF500037	974/735	Pioneer		DG	DG	DG
45D07	AF520565	928/819	Chorismate mutase—*Globodera pallida*	276/2e^{-73}	DG	DG	DG

[a] Clones with an asterisk encode secretory proteins with predicted nuclear localization signals.
[b] Sequences submitted to GenBank with exception of AF273728 and AF006052, which were already in the databases.
[c] Size of the full-length clone with predicted open reading frame (ORF) size; ** indicates not full length.
[d] In situ hybridization of cDNA probes to mRNA specifically within the single dorsal esophageal gland cell (DG) or the two subventral esophageal gland cells (SvG) in preparasitic second-stage juveniles (Pre-J2), parasitic J2 (Par-J2), or parasitic J3, J4, or young adult stages (J3-A) of *Heterodera glycines*.
[e] Novel transcript with no homology to any genes in the public databases.
[f] Not detected.
[g] Percent identity in the amino acid residues of predicted protein.

The SMART oligonucleotide system was used here to obtain the 5′-end of cDNA needed to predict the amino-terminal signal peptide sequence for secretion, while other studies have utilized the conserved splice-leader sequence at the 5′ end of many nematode mRNAs during PCR amplification to create cDNA libraries enriched with full-length cDNAs (Davis et al. 2000; Lambert et al. 1999). Even though approximately 70% of the mRNAs in *C. elegans* contain a splice leader (Blumenthal and Steward 1997), none of the *H. glycines* putative parasitism genes cloned in this study contained a trans-spliced leader, indicating a possible bias against cloning parasitism genes if the splice leader is used as a strategy for cloning gland-expressed genes.

The large proportion (75%) of pioneer putative parasitism genes we obtained here would probably have been overlooked as having potential roles in parasitism if present in a whole nematode EST data set, since data on developmental expression in the esophageal gland cells would not have been available. The suite of pioneer candidate parasitism genes discovered may represent genes unique for parasitism of soybean by *H. glycines*, but this hypothesis is tempered by the limited number of gene sequences currently available in the databases from only a few species of phytonematodes. Nevertheless, BLASTX analysis of the *H. glycines* candidate parasitism genes with 37 *Meloidogyne incognita* candidate parasitism genes (Huang et al. 2003) revealed few similarities, suggesting that the molecular tools used by cyst and root-knot nematodes to parasitize plants may differ considerably. Furthermore, the discovery that only 9% of the putative parasitism genes have homologues in *C. elegans* suggests that this model nematode has limited usefulness in studying nematode adaptations for plant parasitism.

The prediction by PSORT II that 15 *H. glycines* gland-expressed genes encoded secretory proteins targeted for nuclear import makes these putative parasitism genes intriguing candidates for further functional analyses. Nuclear import is primarily mediated by NLSs, which are short polypeptide regions enriched in the basic amino acids arginine and lysine (Hicks and Raikhel 1995; Merkle 2001). Of the three recognized classes of NLSs that have been characterized, nuclear import sequences of the SV40-like and the bipartite classes were present in the predicted nuclear proteins encoded by the putative parasitism genes. Since the predicted nuclear proteins were preceded by an amino-terminal signal peptide for targeting the proteins to the secretory pathway, these proteins are presumably precluded from entering the nucleus of the *H. glycines* esophageal gland cells in which they were synthesized. Indeed, these predicted secretory nuclear proteins are potential candidates for *H. glycines* secretions that could enter the nucleus of a host cell following delivery into the cell through the nematode stylet to regulate soybean root cell phenotype during parasitism. Both types of esophageal gland cells in *H. glycines* expressed genes encoding the predicted secretory nuclear proteins, of which 87% were novel proteins. Other pathogens and parasites are known to release nuclear-targeted proteins into host tissue. For example, pathogenicity proteins containing NLSs have been identified for bacterial plant pathogens, such as *Xanthomonas* species (Yang and Gabriel 1995). Also, *Tri-*

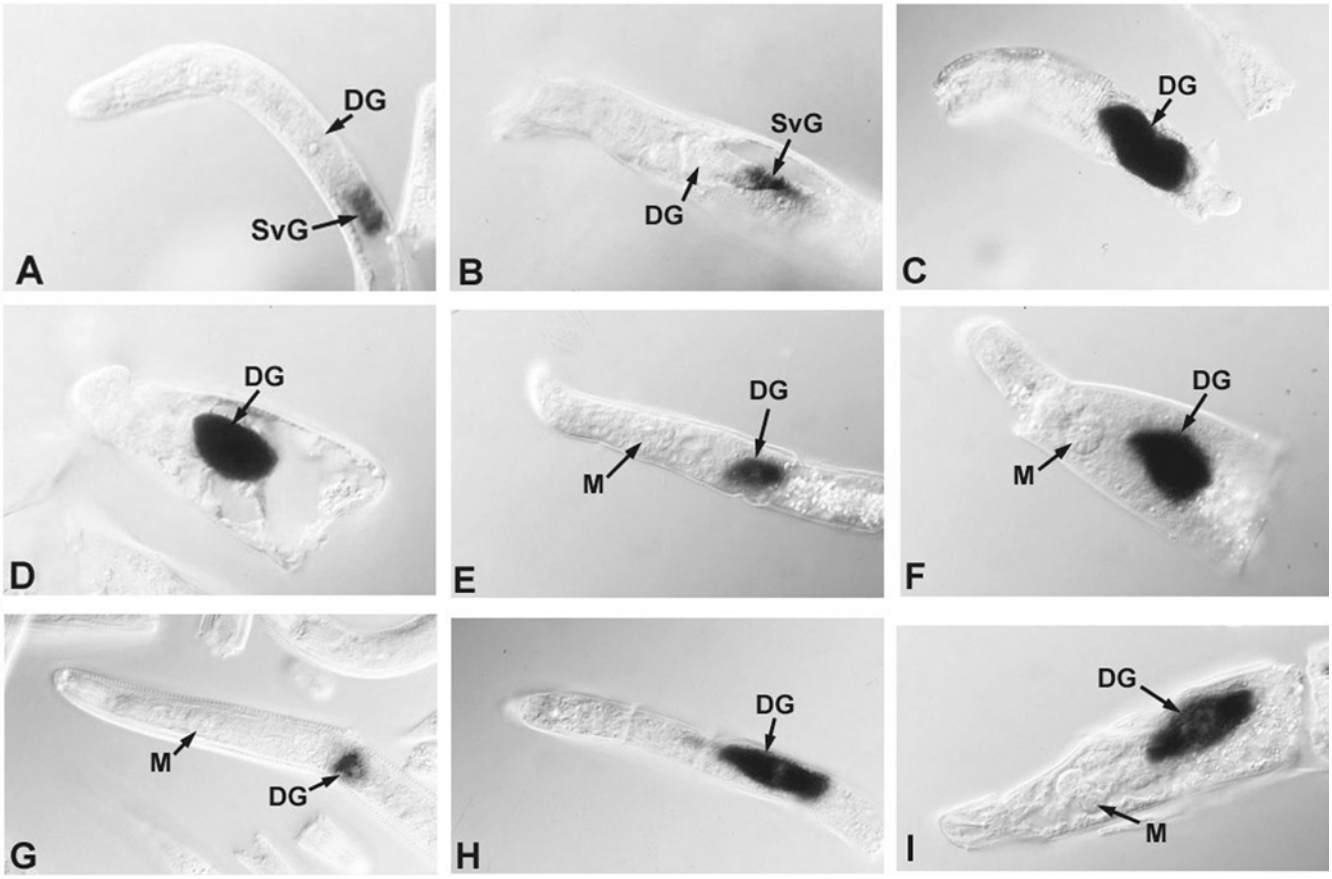

Fig. 1. Hybridization of digoxigenin-labeled antisense cDNA probes (dark staining) of parasitism gene clones to transcripts expressed exclusively within the subventral or dorsal esophageal gland cells of *Heterodera glycines*. Clone 17G06: **A,** migratory parasitic second-stage juvenile and **B,** parasitic third-stage juvenile. Clone *4G12*: **C,** parasitic second-stage juvenile and **D,** parasitic fourth-stage juvenile. Clone 16B09: **E,** parasitic second-stage juvenile and **F,** parasitic fourth-stage juvenile. Clone 28B03: **G,** preparasitic second-stage juvenile, **H,** parasitic second-stage juvenile, and **I,** parasitic fourth-stage juvenile. DG = dorsal gland cell, M = metacorpus, and SvG = subventral gland cells.

chinella spiralis, an animal parasitic nematode, releases an antigen that localizes to nuclei in infected host muscle cells (Yao and Jasmer 1998).

Homologues of the putative parasitism genes in other organisms may provide some insight into the function of the proteins in the *H. glycines*–soybean interaction. The predicted product of clone 3B05 consisted of only a signal peptide and a cellulose-binding domain, which had 34% identity with the cellulose-binding domain of the *H. glycines* endoglucanase HG-ENG-1 (Smant et al. 1998). A cellulose-binding protein with a novel amino-terminal domain has been cloned previously from *M. incognita* (Ding et al. 1998). A recombinant cellulose-binding domain, when expressed in plant tissue, was found to modulate the elongation of different plant cells (Shpigel et al. 1995), indicating a possible role for the novel cellulose-binding protein secreted by *H. glycines* in the development of the large feeding cells (syncytia) in soybean roots. The discovery of three new cellulase genes (6FO6, 13C08, and 26D05) makes a total of six subventral gland cell–synthesized β-1,4-endoglucanases in *H. glycines*. Previous research indicates that *H. glycines* uses these enzymes to hydrolyze the β-1,4 glycosidic bonds of cellulose in the cell walls during penetration and intracellular migration within soybean roots (De Boer et al. 1999; Wang et al. 1999). Interestingly, the predicted β-1,4-endoglucanase of clone 6F06 was similar to the catalytic domain of previously identified *H. glycines* β-1,4-endoglucanases (e.g., AF04421, 96% sequence identity [Gao et al. 2002a]), whereas clones 13C08 and 26D05 encoded distinctly different β-1,4-endoglucanases, with the amino acid identity of the catalytic domains being only 42 to 48% to those of the previously cloned β-1,4-endoglucanases. A different group of cell wall–digesting enzymes is represented by pectate lyases. Clone 17C07 encodes a pectate lyase, which is the second pectinase gene cloned from *H. glycines* (De Boer et al. 2002b). Interestingly, the β-1,4-endoglucanases and pectate lyase from *H. glycines* have their highest similarity outside of plant-parasitic nematodes with genes of prokaryotic microbes, suggesting a potential horizontal gene transfer of these parasitism genes to an ancestor of cyst nematodes (Davis et al. 2000; Smant et al. 1998; Yan et al. 1998).

Another parasitism gene product of *H. glycines* with a distinct prokaryotic signature is chorismate mutase encoded by clone 45B07, which is expressed in the dorsal gland cell. This protein is a key branch-point enzyme in the shikimate pathway leading to the synthesis of phenylalanine and tyrosine. Chorismate mutase was previously cloned from the root-knot nematode, M. javanica, and Globodera pallida, in which it is expressed in the subventral gland cells (Jones et al. 2003; Lambert et al. 1999). Transgenic expression of MJ-CM-1 in soybean hairy roots results in a phenotype of reduced and aborted lateral roots (Doyle and Lambert 2003).

The 4F01-encoded protein resembles a secretory annexin. This finding is interesting because annexins represent a large gene family of calcium-dependent, phospholipid-binding proteins without signal peptides. In *C. elegans*, annexins are expressed in a variety of tissues, including the gland cells of the terminal bulb of the esophagus, where they may have a role in exocytosis (Creutz 1992; Daigle et al. 1999). An annexin was immunolocalized in the amphids and nonglandular tissue in *G. pallida* (Fioretti et al. 2001). The *H. glycines* predicted annexin differed from those nematode annexins by being preceded by a signal peptide and only expressed in the dorsal gland cell. Clone 3D11 encodes a chitinase that is expressed in the subventral glands only in the parasitic stages, precluding a role in egg hatching (Gao et al. 2002b). Interestingly, clone 2B10, which was the most abundant cDNA in the parasitism gene sequences and was previously cloned from *H. glycines* (Wang et al. 2001), contains the C-terminal motif of the bioactive plant peptide CLAVATA3 (Lindsey et al. 2002; Olsen and Skriver 2003). Clone 2A05 encodes the second venom allergen-like protein isolated from *H. glycines* (Gao et al. 2001b). Venom allergen-like proteins constitute a family of secreted proteins found in nematodes with similarity to an extracellular protein from hymenopteran insect venom. The translation products of 3H07 and 4G06 are similar to plant ubiquitin proteins (Callis et al. 1990). However, these *H. glycines* ubiquitins have two unique features. They contain a signal peptide and a novel 19 amino acid extension peptide at the C-terminus. Plant ubiquitins, on the other hand, are not secretory proteins, and their extension proteins are ribosomal proteins (Callis et al. 1990). Since ubiquitin is an abundant protein in a plant cell, the secretion of ubiquitin per se by *H. glycines* would not likely affect the parasitized cell. One possibility is that the secretory ubiquitins of *H. glycines* function as a chaperone for the novel extension peptide that might serve as a peptide signal in the parasitized cell (Lindsey et al. 2002). This possible role is currently being investigated. In addition, *H. glycines* produces another secretory protein involved in ubiquitination, the 8H06 translation product, which resembles an SKP1 protein. SKP1 is a key component of the SCF complex involved in a variety of signal transduction pathways using ubiquitination (Carlos del Pozo and Estelle 1999). Interestingly, the few translation products of the putative parasitism genes with similarity to proteins of known function are either involved in cell wall digestion or suggest involvement in the regulatory machinery of plant cells.

The drastic phenotypic changes of soybean root cells during syncytium formation are the result of *H. glycines*–mediated changes, directly or indirectly, in the developmental program of the parasitized cells (Williamson and Hussey 1996). An understanding of the molecular signaling events in this process will not only provide fundamental knowledge of nematode parasitism and regulation of plant gene expression, it will also suggest vulnerable points in the parasitic process to interfere with for practical applications to limit *H. glycines*–related damage to soybean. Furthermore, intraspecific comparison of the structure of parasitism genes should provide the knowledge to establish a genetic basis for soybean cultivar specificity among *H. glycines* races (Riggs and Schmitt 1988). Beyond the subtle differences in molecular structure of parasitism gene products among *H. glycines* genotypes lies the larger question of how the products of the parasitism genes function and interact with host cell molecules to culminate in successful parasitism. A combination of RNA-mediated interference, inhibitory peptides, and parasitism gene expression in host cells will be required to dissect these functions (Davis et al. 2000). Recently, ingestion of dsRNA by preparasitic second-stage juveniles of cyst nematodes was effective in reducing the corresponding transcripts and development of the nematode on its host (Urwin et al. 2002). The *H. glycines* parasitome discovered here provides a comprehensive template to target functional analyses for dissecting the molecular interactions of *H. glycines* with soybean.

MATERIALS AND METHODS

H. glycines gland cell cDNA amplification by PCR.

Life stages of *H. glycines* representing different timepoints in the parasitic cycle were hand-dissected from infected soybean roots, surface-disinfested, and embedded in 0.7% agarose. The cytoplasm of the esophageal gland cells of 10 *H. glycines* specimens was aspirated into glass micropipettes containing 1 µl of mRNA extraction buffer to minimize nuclease degradation of the mRNA (Gao et al. 2001a) and was transferred into separate microcentrifuge tubes for storage at –80°C until used.

Poly(A)+ RNA from the aspirated gland-cell cytoplasm was isolated on Oligo $(dT)^{25}$ magnetic beads (Dynal, Lake Success, NY, U.S.A.) and eluted with 5:l diethyl pyrocarbonate (DEPC)-treated water at 70°C for 2 min (Gao et al. 2001a). First-strand cDNA synthesis was conducted in 0.5-ml reaction tubes in a 1 µl volume of the following mixture: 4 µl mRNA sample, 0.5 µl 10 FM 3′-RACE cDNA synthesis primer (Clontech Laboratories, Palo Alto, CA, U.S.A.), 0.5 µl 10:M SMART II Oligonucleotide (Clontech), 2.0 µl 5× first strand buffer, 1.0 µl 20 mM DTT, 1.0 µl 10 mM 50X dNTP, 1.0 µl Superscript II (200 units per µl) (Gibco-BRL, Grand Island, NY, U.S.A.). The SMART oligonucleotide system was used to enrich full-length cDNA for subsequent signal peptide analyses. The tubes were incubated at 42°C for 1 h, and 90 µl TE buffer (10 mM Tris, pH 7.6, and 1 mM EDTA) was added. Diluted first-strand reaction solution (10 µl), 2 µl 10 mM dNTP mix, 10 µl TaqPlus long 10x low salt buffer, and 1 µl TaqPlus long (Stratagene, La Jolla, CA, U.S.A.), 2 µl nested universal primer (Clontech) were used in a 100:l volume long-distance (LD) PCR reaction. LD PCR was performed with hot start followed by 24 cycles at 94°C (20 s), 65°C (30 s), and 72°C (7 min). Negative controls of DEPC water were performed at each reaction step above.

Construction of gland-cell cDNA library.

The gland-cell cDNA derived from LD PCR was purified with the QIAquick PCR purification kit (Qiagen, Valencia, CA, U.S.A.), as described previously (Gao et al. 2001a), and ligated into the pGEM-T Easy (Promega, Madison, WI, U.S.A.) vector at a mass ratio of 3:1 (plasmid/cDNA) at 4°C overnight. Ligation products were precipitated with 10 mM glycogen and 100% ethanol, followed by a wash with 70% ethanol. The purified ligation products were electroporated into competent cells of *E. coli* DH10B. White colonies (on blue-white selection) of the gland cell cDNA library were hand-picked (6,130), were transferred to 96-well Microtest III tissue culture plates (Becton Dickinson, Franklin Lakes, NJ, U.S.A.) containing 200 µl 10% glycerol Luria-Bertani medium with ampicillin, and were incubated overnight at 37°C prior to macroarraying (BioRobotics Co., Malden, MA, U.S.A.) onto sterile Hybond-XL nylon membranes (Amersham Pharmacia Biotech, Piscataway, NJ, U.S.A.).

cDNA sequencing and bioinformatics.

The macroarrayed gland-cell library was hybridized with cDNA probes from a LD PCR cDNA library of *H. glycines* intestinal cells, prepared as described above to remove expressed common housekeeping and structural genes. Gel-purified [^{32}P]dCTP labeled intestinal cDNA probes (20 ng) were hybridized to a macroarray of the gland-cell cDNA library and gland cell clones (approximately 4,800) that did not hybridize to intestinal cDNA probes were randomly selected for 5′ end single pass cDNA sequencing using the primer 5′-GGTAACAACGCAGAGTACGCG-3′. Sequences were collected on an ABI 3700 Sequencer (Applied Biosystems, Foster City, CA, U.S.A.). The average sequence read for the single pass sequence was 555 bp. Phred was used to generate base calls and quality scores from the raw chromatograph files from the ABI 3700 Sequencer (Ewing and Green 1998; Ewing et al. 1998). Sequencher (GeneCodes, Ann Arbor, MI, U.S.A.) was used to process (remove low quality sequence from the ends of each read and vector sequence) the individual sequence files, using the Phred base calls and quality scores. Contigs of the high quality sequences (3,711) were assembled using Sequencher parameters of 85% minimum match in at least a 50-bp overlap. Batches of the high quality sequences were analyzed using the BLAST client software, blastcl3, at the National Center Biotechnology Information website. A BLASTX search was performed to determine sequence identity at the protein level, and the results were compiled into a Microsoft Excel-based database with the aid of MuSeqBox (Xing and Brendel 2000). Analysis for an amino-terminal signal peptide sequence for secretion was conducted on all translated processed sequences using SignalP (Nielsen et al. 1997). Transmembrane regions and subcellular localization sites of the deduced proteins were predicted by TMHMM (Krogh et al. 2001) and PSORT II (Nakai and Horton 1999), respectively. The predicted proteins encoded by cDNAs expressed in the gland cells also were analyzed with PSI-BLASTP (Altschul et al. 1997) with novel sequences designated as "pioneers".

Developmental expression of candidate parasitism genes.

Specific forward and reverse primers for each gland-cell cDNA clone predicted to encode a signal peptide were used to synthesize digoxigenin (DIG)-labeled antisense cDNA probes (Boehringer Mannheim, Mannheim, Germany) by asymmetric PCR (Gao et al. 2001a). High-throughput in situ hybridizations were performed using fixed, permeabilized preparasitic juveniles and mixed parasitic stages of *H. glycines* (De Boer et al. 1998; Gao et al. 2001a). cDNA probes that hybridized within the nematode were detected with alkaline phosphatase-conjugated anti-DIG antibody and substrate, and specimens were observed with a compound light microscope (De Boer et al. 1998).

ACKNOWLEDGMENTS

We thank S. Kamoun for help with high throughput signal peptide analyses. Support for this research was provided by the United Soybean Board (Project No. 9214), the Iowa Soybean Promotion Board, the Iowa Agriculture and Home Economics Experiment Station (Project No. 3381), by North Carolina Agricultural Research Service, by Hatch Act and State of Iowa, and by state and Hatch Funds allocated to the Georgia Agricultural Experiment Stations.

LITERATURE CITED

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Blumenthal, T., and Steward, K. 1997. RNA processing and gene structure. Pages 117-146 in: *C. elegans* II. D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, eds. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.

The *C. elegans* Genome Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science 282:2012-2018.

Callis, J., Raasch, J. A., and Vierstra, R. D. 1990. Ubiquitin extension proteins of *Arabidopsis thaliana*. J. Biol. Chem. 265:12486-12493.

Carlos del Pozo, J., and Estelle, M. 1999. Function of the ubiquitin-proteosome pathway in auxin response. Trends Plant Sci. 4:107-112.

Creutz, C. E. 1992. The annexins and exocytosis. Science 258:924-931.

Daigle, S. N., and Creutz, C. E. 1999. Transcription, biochemistry and localization of nematode annexins. J. Cell Sci. 112:1901-1913.

Davis, E. L., and Tylka, G. L. 2000. Soybean cyst nematode disease. The Plant Health Instructor 10:1094. Published online.

Davis, E. L., Hussey, R. S., Baum, T. J., Bakker, J., Schots, A., Rosso, M-N., and Abad, P. 2000. Nematode parasitism genes. Annu. Rev. Phytopathol. 38:365-396.

De Boer, J. M., Yan, Y., Smant, G., Davis, E. L., and Baum, T. J. 1998. In-situ hybridization to messenger RNA in *Heterodera glycines*. J. Nematol. 30:309-12.

De Boer, J. M., Yan, Y., Wang, X., Smant, G., Hussey, R. S., Davis, E. L., and Baum, T. J. 1999. Developmental expression of secretory β-1,4-endoglucanases in the subventral esophageal glands of *Heterodera glycines*. Mol. Plant-Microbe Interact. 12:663-669.

De Boer, J. M., McDermott, J. P., Wang, X., Maier, T., Qui, F., Hussey, R. S., Davis, E. L., and Baum, T. J. 2002a. The use of DNA microarrays for the developmental expression analysis of cDNAs from the esophageal

gland cell region of *Heterodera glycines*. Mol. Plant Pathol. 3:261-270.
De Boer, J. M., McDermott, J. P., Davis, E. L., Hussey, R. S., Smant, G. and Baum T. J. 2002b. Cloning of a putative pectate lyase gene expressed in the subventral esophageal glands of *Heterodera glycines*. J. Nematol. 33:9-11.
Ding, X., Shields, J., Allen, R., and Hussey, R. S. 1998. A secretory cellulose-binding protein cDNA cloned from the root-knot nematode (*Meloidogyne incognita*). Mol. Plant-Microbe Interact. 11:952-959.
Doyle, E. A., and Lambert, K. N. 2003. *Meloidogyne javanica* chorismate mutase 1 alters plant cell development. Mol. Plant-Microbe Interact. 16:123-131.
Ewing, B., and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8:186-194.
Ewing, B., Hillier, L., Wendl, M. C., and Green, P. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res. 8:175-185.
Fioretti, L., Warry, A., Porter, A., Haydock, P., and Curtis, R. 2001. Isolation and localization of an annexin gene (*gp-nex*) from the potato cyst nematode, *Globodera pallida*. Nematology 3:45-54.
Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2001a. Identification of putative parasitism genes expressed in the esophageal gland cells of the soybean cyst nematode, *Heterodera glycines.* Mol. Plant-Microbe Interact. 14:1247-1254.
Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2001b. Molecular characterization and expression of two venom allergen-like protein genes in *Heterodera glycines.* Int. J. Parasitol. 31:1617-1625.
Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2002a. Identification of a new β-1,4-endoglucanase gene expressed in the subventral gland cells of *Heterodera glycines.* J. Nematol. 34:12-15.
Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2002b. Characterisation and developmental expression of a chitinase gene in *Heterodera glycines*. Int. J. Parasitol. 32:1293-1300.
Greenbaum, D., Luscombe, N. M., Jansen, R., Qian, J., and Gerstein, M. 2001. Interrelating different types of genomic data, from proteome to secretome: 'Homing' in on function. Genome Res. 11:1463-68.
Hicks, G. R., and Raikhel, N. V. 1995. Nuclear localization signal binding proteins in higher plant nuclei. Annu. Rev. Cell Dev. Biol. 11:155-88.
Huang, G., Gao, B., Maier, T., Allen, R., Davis, E. L., Baum, T. J., and Hussey, R. S. 2003. A profile of putative parasitism genes expressed in the esophageal gland cells of the root-knot nematode *Meloidogyne incognita.* Mol. Plant-Microbe Interact. 16:5:376-381.
Hussey, R. S. 1989. Disease-inducing secretions of plant-parasitic nematodes. Annu. Rev. Phytopathol. 27:123-141.
Hussey, R. S., and Grundler, F. M. 1998. Nematode parasitism of plants. Pages 213-243 in: Physiology and Biochemistry of Free-Living and Plant Parasitic Nematodes. R. N. Perry and D. J. Wright, eds. CAB International Press, Wallingford, U.K.
Hussey, R. S., Davis, E. L., Baum, T. J. 2002. Secrets in secretions: Genes that control nematode parasitism of plants. Braz. J. Plant Physiol. 14:183-194.
Jones, J. T., Furlanetto, C., Bakker, E., Banks, B., Blok, V., Chen, Q., Phillips, M., and Prior, A. 2003. Characterization of a chorismate mutase from the potato cyst nematode *Globodera pallida*. Mol. Plant Pathol. 4:43-50.
Karrer, E. E., Lincoln, J. E., Hogenhout, S., Bennett, A. B., Bostock, R. M., Martineau, B., Lucas, W. J., Gilchrist, D., and Alexander, D. 1995. In situ isolation of mRNA from individual plant cells: Creation of cell-specific cDNA libraries. Proc. Natl. Acad. Sci. U.S.A. 92:3814-3818.
Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. 305:567-580.
Lambert, K. N., Allen, K. D., and Sussex, L. M. 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. Mol. Plant-Microbe Interact. 12:328-336.
Lindsey, K., Casson, S., and Chilley, P. 2002. Peptides: New signalling molecules in plants. Trends Plant Sci. 7:78-83.
Merkle, T. 2001. Nuclear import and export of proteins in plants: A tool for regulation of signaling. Planta 213:499-517.
Nakai, K., and Horton, P. 1999. PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. 24:34-35.
Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10:1-6.
Olsen, A. N., and Skriver, K. 2003. Ligand mimicry? Plant-parasitic nematode polypeptide with similarity to CLAVATA3. Trends Plant Sci. 8:55-57.
Qin, L., Overmars, H., Helder, J., Popeijus, H., van der Voort, J. R., Groenink, W., van Koert, P., Schots, A., Bakker, J., and Smant, G. 2000. An efficient cDNA-AFLP-based strategy for the identification of putative pathogenicity factors from the potato cyst nematode *Globodera rostochiensis*. Mol. Plant-Microbe Interact. 13:830-836.
Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., Schechter, L. M., Janes, M. D., Buell, C. R., Tang, X. Y., Collmer, A., and Alfano, J. R. 2002. Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. tomato DC3000. Proc. Natl. Acad. Sci. U.S.A. 99:7652-7657.
Ray, C., Abbott, A. G., and Hussey, R. S. 1994. *Trans*-splicing of a *Meloidogyne incognita* mRNA encoding a putative esophageal gland protein. Mol. Biochem. Parasitol. 68:93-101.
Rekhter, M. D., and Chen, J. 2001. Molecular analysis of complex tissues is facilitated by laser capture microdissection: Critical role of upstream tissue processing. Cell Biochem. Biophys. 35:103-113
Riggs, R. D., and Schmitt, D. P. 1988. Complete characterization of race scheme for *Heterodera glycines.* J. Nematol. 20:393-395.
Shpigel, E., Roiz, L., Goren, R., and Shoseyov, O. 1995. Bacterial cellulose-binding domain modulates in vitro elongation of different plant cells. Plant Physiol. 117:1185-1194.
Smant, G., Stokkermans, J. P. W. G., Yan, Y., De Boer, J. M., Baum, T. J., Wang, X., Hussey, R. S., Gommers, F. J., Henrissat, B., Davis, E. L., Helder, J., Shots, A., and Bakker, J. 1998. Endogenous cellulases in animals: Isolation of β-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc. Natl. Acad. Sci. U.S.A. 95:4906-4911.
Urwin, P. E., Lilley, C. J., and Atkinson, H. J. 2002. Ingestion of double-stranded RNA by preparasitic juvenile cyst nematodes leads to RNA interference. Mol. Plant-Microbe Interact. 15:747-752.
Xing, L., and Brendel, V. 2000. MuSeqBox: A program for multi-query sequence BLAST output examination. Bioinformatics 17:744-745.
Wang, X., Meyers, D., Yan, Y., Baum, T., Smant, G., Hussey, R., and Davis, E. 1999. In planta localization of a β-1,4-endoglucanase secreted by *Heterodera glycines*. Mol. Plant-Microbe Interact 12:64-67.
Wang, X., Allen, R., Ding, X., Goellner, M., Maier, T., de Boer, J. M., Baum, T. J., Hussey, R. S., and Davis, E. L. 2001. Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. Mol. Plant-Microbe Interact. 14: 536-544.
Williamson, V. M., and Hussey, R. S. 1996. Nematode pathogenesis and resistance in plants. Plant Cell 8:1735-1745.
Yan, Y., Smant, G., Stokkermans, J., Qin, L., Helder, J., Baum, T., Schots, A., and Davis, E. 1998. Genomic organization of four β-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. Gene 220:61-70.
Yang, Y., and Gabriel, D. W. 1995. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. Mol. Plant-Microbe Interact. 8:627-631.
Yao, C., and Jasmer, D. P. 1998. Nuclear antigens in *Trichnella spiralis* infected muscle cells: Nuclear extraction, compartmentalization, and complex formation. Mol. Biochem. Parasitol. 92:207-218.

AUTHOR-RECOMMENDED INTERNET RESOURCES

National Center Biotechnology Information BLAST server website: www.ncbi.nlm.nih.gov/blast/index.html
Center for Biological Sequence Analysis (CBS) SignalP server website: www.cbs.dtu.dk/services/SignalP/index.html
CBS transmembrane helices in proteins prediction (TMHMM) website: http://www.cbs.dtu.dk/services/TMHMM/